

ABSTRACT OF THE DISCLOSURE
A HIGHLY SCALABLE AND HIGHLY AVAILABLE
CLUSTER SYSTEM MANAGEMENT SCHEME

5 A cluster system is treated as a set of resource
groups, each resource group including an highly available
application and the resources upon which it depends. A
resource group may have between 2 and M data processing
systems, where M is small relative to the cluster size N of
10 the total cluster. Configuration and status information for
the resource group is fully replicated only on those data
processing systems which are members of the resource group.
A configuration object/database record for the resource
group has an associated owner list identifying the data
processing systems which are members of the resource group
15 and which may therefore manage the application. A data
processing system may belong to more than one resource
group, however, and configuration and status information for
the data processing system is replicated to each data
processing system which could be affected by failure of the
20 subject data processing system--that is, any data processing
system which belongs to at least one resource group also
containing the subject data processing system. The partial
replication scheme of the present invention allows resource
groups to run in parallel, reduces the cost of data
25 replication and access, is highly scalable and applicable to
very large clusters, and provides better performance after a
catastrophe such as a network partition.